THE CRITICAL PERIOD HYPOTHESIS FROM AN ENGLISH AS A FOREIGN LANGUAGE PERSPECTIVE

by

Dominique Samantha Pantophlet

A thesis submitted to the faculty of
The University of Utah
in partial fulfilment of the requirements for the degree of

Master of Arts

Department of Linguistics

The University of Utah

May 2010

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

Dominique Samantha Pantophlet

This thesis has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

Rachel Hayes-Harb

Rubio

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the thesis of Dominique Samantha Pantophlet in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Chair: Supervisory Committee

Approved for the Major Department

Chair/Dean

Approved for the Graduate Council

Charles A. Wight
Dean of The Graduate School

ABSTRACT

The Critical Period Hypothesis (CPH), originally proposed in Lenneberg (1967), states that there is a maturational period of time during which language acquisition can take place. The CPH is often cited to account for differences in success between early and late language learners. Based on Lenneberg's hypothesis, other researchers have generated additional hypotheses, all of which consider a maturational period of time during which language acquisition can take place. This thesis investigates the critical period hypothesis for second language (L2) pronunciation from an English as a foreign language perspective. It investigates whether it is possible for late learners of English to achieve a native-like pronunciation regardless of maturational constraints.

This thesis also investigates whether the variable 'exposure to the target language' significantly influences the ability of late L2 learners to obtain native like pronunciation. In order to investigate whether the variable exposure to the target language significantly influences the ability of late L2 learners to obtain a native-like pronunciation, two groups of non-native speakers, differing in the amount of exposure they received to the target language in their secondary education, were included in this study alongside a native speaker control group.

Three speech samples were collected for each speaker-- a word list, a paragraph, and an answer to an open-ended question. After three speech samples were collected for each subject, 21 linguistically naive native speakers of English evaluated the sound samples based on native likeness. After these scores were obtained, means were calculated for each individual speaker, in order to determine their performance, and for speaker groups, in order to determine intergroup comparison.

Results indicate that there were no late L2 learners of English who had obtained a native-like pronunciation. This could provide evidence in favor of a critical period for second language pronunciation. Results also indicate that there was a significant difference in mean scores between the non-native speaker groups, with an advantage for students in bilingual Dutch-English programs over those in monolingual Dutch programs, suggesting a significance for the variable 'exposure to the target language' in determining second language pronunciation in late learners. Thus, exposure to the target language influences the acquisition of second language pronunciation, possibly, alongside a developmental critical period.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my supervisory committee, consisting of my Chair, Dr. MaryAnn Christison, Dr. Rachel Hayes-Harb, and Dr. Fernando Rubio, for their supervision and support throughout this project.

I would also like to extend my gratefulness to the schools and subjects in The Netherlands and Utah who participated in this study. Without their willingness to cooperate and their efforts to provide me with their feedback in an efficient and timely manner, I would not have been able to write this thesis. I hope that their efforts will benefit research on foreign language learning and that they may positively influence secondary education in The Netherlands.

CHAPTER 1

INTRODUCTION

This thesis aims to investigate the Critical Period Hypothesis (CPH) for second language pronunciation from an English as a foreign language perspective by researching the possibility that late learners of English are able to achieve a native-like pronunciation in the target language. Laypersons often have the perception that adults or late learners of second languages sound like non-native speakers when producing a second language. The perception also persists that early learners or children seem to be able to attain a native-like pronunciation of a second language. Even though late learners may have a cognitive advantage over early learners, in such a way that they may have developed increased abstract thinking skills over time that could help them in processing second language input, this cognitive advantage does not seem to provide an advantage for native-like pronunciation. The question that researchers ask, is the following: Why do early learners often seem to be able to attain a native-like second language pronunciation with more success than late learners?

Research from a variety of disciplines suggests that one reason for the difference in success between early and late second language learners relative to native like pronunciation can be found in a biologically based critical period for second language acquisition. Hypotheses about such a critical period vary considerably, ranging from considering one absolute time span to determine successful language acquisition as a whole (Lenneberg,

1967), to multiple possible time spans that determine the ease with which specific skills can be acquired, such as for pronunciation (Flege, Munro, and MacKay, 1995), morphology, and syntax (Coppieters, 1987), but all are based on the assumption that there is a maturational period or time span during which language acquisition can (easily) take place (Lenneberg, 1967). The premise for all of these hypotheses is that, if language acquisition commences after a certain time period (e.g., puberty) has passed, learners will not be able to achieve a native-like proficiency in the target language.

This study investigates layperson perceptions of non-native speaker (NNS) speech and whether it is possible to find late learners who have obtained a native-like pronunciation. The reason that this question is important is that, if it is possible for late learners of a second language to acquire the pronunciation of such language to a level that is similar to that of native speakers, it at least calls into question the strength of the CPH if not the existence of the hypothesis itself.

Although studies such as Bongaerts, Planken, and Schils (1995) and Ioup (1994) have been conducted to investigate whether it is possible to find late learners who have acquired a native-like pronunciation of a second language, they all focus on preidentified, non-native, and highly advanced speakers. The subjects in the above studies were indeed late learners of English, as they had not commenced their L2 acquisition until the age of 12, but they were individually selected to participate in the study based on their high level of English proficiency. The present study, on the other hand, used an opportunity sample so as to investigate the ability of an average group of late English language learners to obtain a native-like pronunciation of the target language. In contrast to the subjects of the present study, most of the subjects in Bongaerts et al. (1995) had lived abroad for a number of years or were linguistically trained English professors. As such, these researchers demonstrate that

it is indeed possible to find exceptional individuals, but they fail to investigate whether the variable of exposure to the target language is of as much influence in large groups of subject that are not preidentified as it is on these exceptional individuals. In other words, they are not able to establish whether exposure to a target language positively influences second language acquisition for everyone, or just for a select few.

In addition, most research on the topic of late non-native second language learners, such as Flege, Birdsong, Bialistok, Mack, Sung, and Tsukada (2005) and Flege et al. (1995), is conducted in target language settings in which learners are exposed to the target language in a variety of ways and contexts. Consequently, it becomes difficult, if not impossible, to truly consider the amount to which they are exposed to the target language.

This thesis, consequently aims to fill this gap by using groups of students who differ greatly in the amount of target language exposure they have received (i.e., whether they received monolingual or bilingual education). In addition, subjects have not been preidentified as L1 speakers of Dutch who are advanced L2 speakers of English. As the subjects in this study are not part of the target culture, since they live in The Netherlands, their exposure to the target language is highly limited to the instructional domain, which allows for the variable 'exposure to the target language' to be considered. Exposure to the target language outside of the instructional domain is relatively similar for all speakers as they are part of the same culture and reside in the same geographical area. It is, however, not entirely possible to fully extract the influence of target language instruction from that of target language exposure outside of the instructional domain as objective measures of target language exposure outside of the instructional domain are difficult, if not impossible, to obtain, which presents a limitation on this study. Due to the spread of English as an international language of communication, media, and technology, however, it is questionable

whether it is possible to find a context in which these two types of English language input can be measured independently from one another.

In order to test the influence of a critical period on the acquisition of second language pronunciation in late learners, 12 L1 speakers of Dutch and 6 native speakers of English participated in three tasks. The 12 non-native speakers of English had all received 6 years of English language instruction during their secondary education at the same school and none of them had commenced their English language acquisition through instruction before the age of 12. All speakers, both native and non-native, had obtained a total of 12 years of public education.

A rater group consisting of 21 native speakers of English rated the obtained sound samples. Untrained native speaker raters were used in order to obtain judgments about the pronunciation of native and non-native speakers of English, as the judgments of untrained raters were considered to be typical of the perceptions of untrained, or laypersons. Training raters, as such, would not achieve the intended goal, because it would eliminate the possibility for raters to use their personal judgments, and biases as native speakers. Instead, trained raters would apply a taught evaluation procedure. Raters were asked to rate the samples based on the 'native likeness' of the sample of speech on a 5-point Likert scale, because the ability to distinguish the pronunciation of native speakers from that of non-native speaker subjects was thought to demonstrate the influence of a critical period on second language pronunciation. In other words, if significant differences in the pronunciation of native and non-native speakers of English cannot be found in the mixed sound samples, and native and non-native speakers receive similar scores, the successful attainment of second language pronunciation in late learners is likely not restricted by maturational constraints.

In addition to examining the strength of a critical period for the acquisition of second language pronunciation in late learners, this thesis also aims to investigate whether the variable 'exposure to the target language' is of significance in determining second language pronunciation in late learners. In this study, exposure to a target language is defined in terms of two types of instruction for Dutch L1 speakers in the instructional domain. The first type of instruction is monolingual education. In this type of program students receive an average of two hours of English language instruction per week for a duration of 6 years. The second type of instruction is bilingual education in which students receive an average of 14 hours of target language instruction for a duration of 6 years. In order to investigate the effect of the variable exposure to the target language on the acquisition of second language pronunciation in late learners, then, two groups of non-native speakers, differing in the amount of exposure they have received to the target language, are included in this study alongside a native speaker control group.

Although it may seem natural to assume that more exposure to the target language over an extended period of time would result in improved pronunciation, from an innatist perspective, this may not necessarily the case. According to cognitive theories of second language acquisition that adhere to the idea of a critical period, differing amounts of exposure to a target language may not function as the most significant factor in determining the possible end result for second language acquisition (Schachter, 1988). The level of attainment of native-like pronunciation skills in the target language is considered to be governed by biological factors and, as such, is considered a manifestation of innateness rather than one of exposure or practice (Lenneberg, 1967). Consequently, if we consider only biological factors, the two non-native subject groups in this study would be expected to

perform similarly to one another since as non-native speakers with the same L1, both groups would have the same access to whatever innate capacity governs L2 acquisition.

The following research questions motivate this study:

1. Are Dutch L1 speakers in this study who are late learners of English able to achieve a native-like pronunciation of English as based on comparison scores between samples of their speaking and those of L1 speakers of English?
2. Does exposure to the target language make a difference in the level of attainment of pronunciation in Dutch L1 speakers in this study who are late learners of English? In other words, is there a difference in English language pronunciation between students in monolingual education programs in this study and those in bilingual education programs in this study in The Netherlands as based on comparison scores between samples of speaking scores for each group?

Alongside these research questions, the following hypotheses are developed:

Hypotheses for research question one:

H0 There will be no difference in the level of attainment of a native-like pronunciation of English between students in bilingual programs and those in monolingual programs in The Netherlands as based on comparison scores between samples of their pronunciation.

H1. Students in bilingual education programs in The Netherlands will have obtained a more native-like pronunciation in English than students in monolingual education programs in The Netherlands as based on comparison scores between samples of their pronunciation.

Hypotheses for research question two:

H0. There will be Dutch L1 speakers in this study who acquired English as a foreign language after the proposed critical period has passed who will have obtained a native-like pronunciation in English as based on comparison scores between samples of their speaking and those of L1 speakers of English.

H1. There will be no Dutch L1 speakers in this study who acquired English as a foreign language after the proposed critical period has passed who will have obtained a native-like pronunciation in English as based on comparison scores between samples of their speaking and those of L1 speakers of English.

A review of the literature that was used to formulate these research questions can be found in Chapter 2, as well as a background to the situation in which the subjects in this study acquired their first and second language(s). The methods that were used to investigate the research questions mentioned above are carefully detailed in Chapter 3.

The proposed research questions and hypotheses are evaluated in Chapter 4 and Chapter 5 of this thesis in order to consider the strength of a possible critical period for the development of second language pronunciation and its implication on research in the field of second language acquisition.

CHAPTER 2

BACKGROUND

2.1 Review of the Literature

2.1.1 An Introduction to the Critical Period Hypothesis

The CPH for language acquisition has undergone many changes over time, but was first introduced by Lenneberg (1967). In this hypothesis the phenomenon was described as a neurologically based period, ending around the onset of puberty, which "[...] termination seems to be related to a loss of adaptability and inability for reorganization in the brain, particularly with respect to the topographical extent of neurophysiological processes." (p.179). As such, Lenneberg argued that is impossible for a person to learn a second language (L2) with a native-like proficiency after the start of puberty. Based on Lenneberg's hypothesis, other researchers have generated new hypotheses, all of which consider a maturational period of time during which language acquisition can take place. Examples of such generated hypotheses include, for example, variants containing separate critical periods for specific language skills, such as proposed by Flege et al. (1995). Much research has been dedicated to this topic and evidence from a variety of disciplines has been proposed to account for the observed differences in language acquisition between children and adults, such as Penfield and Roberts (1959), Lenneberg (1967), and Flege et al. (1995).

Based on observational arguments (i.e., generally lay perceptions) and those from the field of neurobiology, the Lenneberg (1967) version of the CPH states that the process of

lateralization of language functions is responsible for the difference in language attainment between children and adults. During the process of lateralization, specific areas of the brain, such as Wernicke's and Broca's areas, become specialized in specific language functions. Once set, these language functions cannot be performed by alternate areas of the brain. Lateralization is generally believed to take place around the onset of puberty (Lenneberg 1967), so that, if damage to areas of the brain takes place after puberty has set in, other areas cannot take over the language functions and knowledge and/or skills will be permanently lost. In contrast, other researchers, such as Penfield and Roberts (1959), suggest that if children experience damage to language-governing areas of the brain before lateralization is completed, linguistic functions are generally taken over by different areas of the brain so as to maintain the knowledge and skills available.

More recent evidence from the field of neurobiology considers the process of myelination as a possible underlying cause of the discrepancy in language attainment between early and late learners. Myelination, according to Pulvermüller and Schumann (1994), takes place when glial cells cover the axons of neurons with a fatty sheath, making it easier for the axon to receive nutrition and to function more efficiently. As a result of this process, however, the neural connections become so efficient that it becomes more difficult for the axons to make new connections with other available dendrites, a process that is essential for both first and second language acquisition. Although the precise relationship between myelination and an increased difficulty for neurons to make connections has not yet been determined, it seems to be important. An increase in myelination has also been observed to co-occur with a decrease in plasticity of the brain. The plasticity in the areas of the brain that are concerned with specific language functions is gradually reduced until around puberty. After puberty, plasticity is reduced and is thought to negatively correlate

with age, so that it is more difficult for adult brains to respond as quickly as the brains of early language learners (Pulvermüller & Schumann, 1994).

A third source of evidence for the existence of a critical period for language acquisition can be found in the development of language in feral children. Cases such as Genie (Curtiss, 1977), and Victor (Itard, 1962) suggest that in late first language acquisition, it is extremely difficult, if not impossible, for language learners to reach a native-like attainment. Research on late first language acquisition in hearing adults is rare and generally concerns circumstances in which subjects are deprived of more than just language attainment. In the case of Genie, for example, the subject of research was exposed to extreme deprivation and abuse. Her failure to fully attain a native-like proficiency in a first language after close practice is often used as evidence for the idea of a critical period; however, given her background it could be argued that there may be different causal factors that account for her inability to obtain a native-like first language proficiency. Instead of considering a critical period as the cause for her inability to attain English, factors such as severe emotional trauma or developmental difficulties should be considered as well in finding an explanation for Genie's limited level of attainment in first language acquisition. Due to the nature of research on late first language acquisition in hearing adults, results cannot safely be generalized to other situations and, as such, cannot in themselves prove the existence of a critical period for language acquisition.

As evidence from the language attainment of feral children suggests, first language acquisition can follow particular patterns that cannot be found in second language acquisition, as second language learners have necessarily obtained a first language prior to or during their immersion in a second language (as in the case of bilingualism in early learners). As such, evidence from situations of language deprivation used in the formation of a critical

period hypothesis for first language acquisition cannot be used to postulate theories on the development of second language acquisition as they are build upon contextual factors that are necessarily impossible in situations of second language acquisition.

Another source of evidence for a critical period can be found in studies on deaf individuals. These studies provide more tangible evidence for the existence of a critical period. Such studies have, for example, shown that late acquisition of sign language generally takes place more effectively if students have already been exposed to a spoken language, prior to their loss of hearing, at an early age. Lenneberg (1967) postulates in more detail that if children are exposed to sign language instruction for a minimum of one year before losing their hearing, they can attain a high level of sign language proficiency much more successfully than if they are first exposed to it in adulthood. Mayberry (1993), similarly, investigates the difference in language acquisition between late first language learners and late second language learners of American Sign Language (ASL). She finds that after using ASL for an average of 50 years, late second language learners of ASL generally outperform late first language learners of ASL on a number of different tasks. When compared to early first language learners of ASL, late second language learners perform remarkably similar, especially on measures of syntax, suggesting that it is still possible for late learners to acquire a native-like ability in a second language if given enough practice.

Based on this evidence, the CPH- in its many forms- can be argued to be of interest for the field of Second Language Acquisition (SLA), as it attempts to account for the complex nature of the level of L2 attainment. Verification of the CPH would provide further evidence for the theory of Universal Grammar and the question of access in second language acquisition (Ellis, 1994). Counterevidence for the existence of a critical period consequently, could suggest limitations on access to Universal Grammar in SLA. If the CPH were to be

seriously questioned, theories of Universal Grammar and the innate capacity for language learning would have to be reconsidered and different theories on language development would have to be posed to account for the difference in language acquisition between adults and children.

Although considerable research has been dedicated to finding evidence both for and against the existence of a critical period, no clear-cut answer has been finalized. At present, according to Long (2007), there are over 100 empirical studies on the topic of maturational constraints on language acquisition, but no consensus has been reached on the "[...]existence, scope, and timing of putative maturational constraints on the human language learning capacity as well as on implications for practice." (p. 46). Further evidence is needed to conclusively eliminate a number of hypotheses, so as to enable the evidence-based formation of a single theory on language acquisition.

2.1.2 Types of Hypotheses

It should be noted, here, that different types of critical period hypotheses are formulated for both first and second language acquisition, but that, for the purpose of this thesis, only those types of hypotheses for second language acquisition will be considered throughout the rest of this thesis. A number of different hypotheses concerning maturational constraints on language learning have been proposed. Although the content of specific hypotheses differs with regard to their scope, or the skills that are affected by maturational constraints, and their timing, or age at which maturational constraints are most potent (Long, 2007), discussing each of them is beyond the scope of this project. They can, however, be divided into the following types of hypotheses:

2.1.2.1 The Critical Period Hypothesis in its Strong Form

The term 'Critical Period Hypothesis' is generally used to refer to a type of hypothesis that argue for a specific cut-off point in age after which it is impossible for language learners to obtain a native-like proficiency in a second language. Also referred to as the 'Maturational State Hypothesis' (Long, 2007), theories in this category assert a "[…] genetically inherited, language acquisition capacity [which] operates only within a genetically determined period, and no later, whether or not exercised during that period." (p. 48). Proponents of this type of hypothesis, thus, believe that language learners can only obtain full proficiency in a language within a certain, specified period, and that after this period, language learners will either not be able to obtain a language at all, or will, at best, have a detectable foreign accent in the language at question.

The Critical Period Hypothesis, or 'Maturational State Hypothesis', for second language acquisition, specifically argues that regardless of the onset of first language acquisition, fluency in a second language cannot be obtained to a native-like extent if it takes place after a specified age. Although theories differ as to the precise age at which this ability to acquire a second language ends, they assert that adults will necessarily be less successful acquirers of second languages than children. This hypothesis supports the common sense belief of 'earlier is better' for the purpose of second language learning and, as such, have gained much support and interest.

Several studies, however, such as Bongaerts et al. (1995), have measured native-like levels of language proficiency by non-native speakers who have acquired a second language after the close of proposed critical periods. These studies provide evidence against the Critical Period Hypothesis in its strong form. Particular examples of such studies will be discussed in

the section 2.1.3, along with other evidence that may disprove different types of critical period hypotheses.

2.1.2.2 The Exercising Hypothesis

A second type of critical period hypothesis can be found in the so-called 'Exercising Hypothesis' (Long, 2007). This type of hypothesis suggests that "[...]once used, or exercised, within the genetically determined period, the language acquisition capacity is available, undiminished, for life." (p. 47). As such, individuals who start the acquisition of a first language during a developmentally based, predetermined period, will benefit from their efforts any time after this predetermined period has ended. This type of hypothesis asserts that late first language learners will necessarily perform worse than both early first language learners, early second language learners, and late second language learners.

Based on this hypothesis, adult and child language learners have the same potential in acquiring a second language if the specific language capacity is exercised prior to a genetically predetermined period. As such, this type of hypothesis fails to account for the difference in second language attainment that is often witnessed between children and adults as it does not leave any possibility for these differences to be observed. According to this hypothesis, there should not be an observable difference in second language attainment between early and late learners as both types of learners have exercised their language capacity for their first language, so that it should be available to the same extent for both types of learners. As pointed out in Krashen, Long, and Scarcella (1979), however, adults may have an initial advantage in cognitive development over children, but younger learners generally outperform adult learners in second language acquisition where ultimate attainment is concerned. According to the exercising hypothesis, this should not be possible as both early and late second language learners have the same access to their language capacity. The

'Exercising Hypothesis' has, consequently, often been rejected based on its inability to account for specific differences in the second language acquisition of children and adults.

2.1.2.3 The Sensitive Period Hypothesis

The 'Sensitive Period Hypothesis', which argues for the existence of a sensitive period, as opposed to a critical period, to account for language acquisition can be considered a third type of critical period hypothesis. The Sensitive Period Hypothesis, thus, argues that there is a single developmentally-based predetermined period during which it is much easier for humans to attain proficiency in a second language (Long, 2007) . This type of hypothesis, thus, does not argue that it is absolutely impossible for an individual to attain native-like or advanced proficiency in a language when commencing their language acquisition after the sensitive period is ended, it merely states that when started during a sensitive period, individuals can attain language proficiency with greater ease due to maturational constraints.

Although specific theories differ as to the timing of this sensitive period, proponents of the Sensitive Period Hypothesis assert that it is necessarily more difficult for adults to obtain advanced proficiency in a language than it is for younger learners, which, again, supports the mainstream idea of 'earlier is better', for the purpose of language learning, though in this case, later is not necessarily impossible.

Evidence against this type of CPH is hard to establish, considering the fact that the Sensitive Period Hypothesis does no argue against the possibility of language acquisition after a sensitive period has ended, but instead mainly suggest that it is necessarily more difficult for late learners to acquire language proficiency than it is for early learners. Unless large numbers of late language learners are found who have obtained native-like proficiency in a language while having commenced their language acquisition after the proposed sensitive period has ended while, at the same time, large groups of young language learners

are identified who have failed in the native-like acquisition of a second language, it seems as though this type of critical period hypothesis cannot be challenged. Although some studies, some of which will be discussed below, have aimed to do precisely this, it is difficult to establish whether their results truly disprove the Sensitive Period Hypothesis per se or whether they simply point at the necessity for specific theories on the Sensitive Period Hypothesis to reconsider their views on the matter of timing.

2.1.2.4 The Multiple Sensitive Periods Hypothesis

In addition to the Sensitive Period Hypothesis discussed above, there is also a type of CPH that includes multiple sensitive periods for language acquisition. The Multiple Sensitive Periods Hypothesis generally argue for a domain-based separation of sensitive periods during which particular components of language acquisition can occur with ease. Although theories differ as to the timing of each sensitive period, most distinguish separate periods for at least the following language domains: phonology, morphology, and syntax. The Multiple Sensitive Periods Hypothesis, thus, argues that there is a developmentally-based predetermined period of time during which language acquisition in each of the domains can take place with ease and that these periods of time can be different for each component. The sensitive period for morphology, could, thus, have an earlier onset than that of syntax, and the onset for phonology could, likewise, precede that of morphology (Long, 2007).

Although the Multiple Sensitive Periods Hypothesis distinguishes multiple periods during which particular language components can be obtained, it does not assert that it is impossible for individuals to attain fluency in these language components if language acquisition is commenced after the proposed sensitive period, it merely suggests that acquisition is more likely to result in a native-like proficiency if it is started before the end of the sensitive period in question.

Even though particular theories on the Multiple Sensitive Periods Hypothesis differ as to the timing of specific sensitive periods, they postulate that it is necessarily more difficult for late learners to acquire language proficiency than it is for early learners, as early learners engage in the process of L2 acquisition during this predetermined maturational period and late learners do not. This supports the mainstream idea of 'earlier is better' for the purposes of language learning, though, in this case, later is not necessarily impossible.

Evidence against the Multiple Sensitive Period hypothesis, thus, is hard to establish, considering the fact that the existence of multiple sensitive periods does not argue against the possibility of language acquisition after a sensitive period has ended, but mainly attempts to account for differences in ease of acquisition between early and late learners. Based on a predetermined maturational period, it is necessarily more difficult for late learners to attain a native-like proficiency in a second language than it is for early learners as early learners engage in the process of L2 acquisition during this predetermined maturational period and late learners do not.

Similar to the case of the 'simple' Sensitive Period hypothesis, it is extremely difficult to disprove the existence of multiple sensitive periods for language acquisition as 'ease of acquisition' can be defined in many different ways, depending on factors such as, for example, speed or success. Unless large numbers of late language learners are found that have obtained native-like proficiency in specific components of a language while having commenced their language acquisition after the proposed sensitive period for the component in question has ended while, at the same time, large groups of young language learners are identified who have failed in the native-like acquisition of such second language components, it seems impossible to argue that there could not be any developmentally predetermined periods during which it is easier for individuals to obtain specific components

of a second language. Although some studies, some of which will be discussed in the next section, have aimed to do precisely this, it is difficult to establish whether their results provide counter evidence to disprove the existence of multiple sensitive periods per se or whether they simply point at the necessity for specific theories on the Sensitive Period Hypothesis to reconsider their views on the matter of timing.

2.1.3 Evidence Against the Critical Period Hypothesis

One way to provide counter evidence for the existence of the types of 'critical' period hypothesis mentioned above would be to find a large number of non-native speakers of a language who have commenced to acquire that language after the critical period has ended and who have attained a native-like proficiency either in a specific component of a second language or in the second language overall. Although this would technically still not disprove the Sensitive Period Hypothesis per se, it, at least, encourages one to reconsider the strength of the effects of a critical period on second language attainment. As Long (2007) points out, " because a native-like accent in a second language or dialect seems so difficult to achieve, more than any other linguistics domain, demonstration of a second language phonology would undermine claims for sensitive periods in SLA." (p. 63).

Although a number of studies have sought to demonstrate native-like second language attainment of late learners for the domain of phonology, there are a number of methodological issues in these studies that need to be considered. Flege et al. (2005), for example, evaluate the influence of age and length of residence on the degree of foreign accent in second language attainment by using a method of evaluation through which subjects are not necessarily tested on their comprehension or language proficiency per se. In their model, subjects are asked to reproduce a given sentence, which does not require any understanding of the speech sample to be produced, and as such indicates a degree of

success in imitation rather than level of L2 proficiency in terms of language proficiency as defined in Cummins and Swain's model of Communicative Competence. Though Flege et al. argue that they inserted a question in between the initial sentence and the moment of repetition to " prevent direct imitation from sensory memory", they fail to take into consideration that the practice session may affect subjects' responses to the later recorded information. Results obtained in this fashion, thus, do not necessarily indicate the degree of foreign accent by speakers of different ages in real life, but simply measure the degree of foreign accent based on selected speech samples in a highly controlled environment. In addition, the subjects in this study resided in the target culture, and, as such are likely exposed to differing degrees of target language input based on personal characteristics. Though measures of self-reporting were used, true differences in exposure to the target language could easily have influenced the effects of length of residence on the language proficiency of the subjects in this study and, as such, could be argued to undermine any evidence against the Critical Period Hypothesis that is suggested on this basis.

Bongaerts et al. (1995), on the other hand, employed four distinctive measures with a decreasing level of speaker control on production in order to investigate the language ability of Dutch late learners of English. Subjects in this study were first asked to discuss a recent vacation for about three minutes, through which spontaneous speech was elicited. Then, subjects were asked to read a short text, followed by 10 short sentences, and finally a wordlist that comprised of 25 words. Each participant could perform each task only once, in order to elicit genuine responses and to eliminate opportunities for self-correction. Speech samples were then rated by native-speakers of English on a 5-point scale based on the degree of accentedness.

Due to rater bias on the dialect of English that was spoken by a number of the subjects in this study, Bongaerts et al. chose to conduct a smaller scale follow up study in which the native speaker control group consisted of speakers of a variety of English that is considered to be more "neutral" than the dialect of English of the control that was used in the initial study. Subjects in the follow up study were asked to read six sentences out loud for a total of three times per sentence. Each sentence was then included in a stimuli pool that was presented to native-speaker judges, who rated them on the 5-point scale that was used in the initial study. For matters of analysis, non-native speakers of English were considered to have produced native-like levels of proficiency if their mean rating score fell within two standard deviations of the mean native-speaker rating score and based on this, five non-native speakers of English were considered to have attained native-like levels of second language proficiency. Bongaerts et al., thus, found that it is possible for late second language learners to obtain a native-like proficiency, which argues against the CPH in its strong forms, though their findings may not necessarily be generalized to different situations. Not only are Dutch and English part of the same language family, and therefore, could be argued to be linguistically similar, the five native-like, non-native subjects in this study were also preidentified as advanced speakers of English and, as such, may not represent Dutch learners of English in general.

In addition to the methodological considerations mentioned above, most research on the CPH fails to define the concept of 'ultimate attainment' and does not consider effects of the variable 'exposure to the target language'. The Critical Period Hypothesis generally state that it is either impossible or unlikely for late language learners to obtain a native-like proficiency in a second language, and studies that aim to prove this hypothesis include subjects of different age groups in order to demonstrate this principle. They do not

generally, however, consider the possibility for subjects to continue to improve their second language proficiency after the study has ended and, as such, fail to recognize that, perhaps, when given sufficient exposure to the target language, subject groups may well outperform others over time.

As mentioned previously, most research on the CPH for second language acquisition has been conducted in second language cultures in which subjects acquire a second language in the target language culture and therefore receive high exposure to the TL in different settings. As such, most studies have failed to take differences in the amount of exposure to the target language into consideration or have based their analysis on estimations and methods of self-reporting.

Long (2007) similarly points out that many studies that are aimed to disprove the CPH employ tests or stimuli that are not of a sufficient level of difficulty for non-native speakers in term of L2 proficiency. If stimuli are included that even low level non-native speakers can understand or produce, possibly due to similarities between the target language and their L1, true differences in ability levels between native speakers and non-native speakers cannot be ascertained.

In order to address these issues of differing amounts of exposure to the target language and linguistically undemanding tests, similar to van Boxtel, and Coppen (2005), the present study employs three measures with an increasing level of speaker control on production so as to include both language and meaning instead of "meaningless streams of sound"(Long, 2007, p. 63). In addition, subjects in the study were not preidentified as advanced speakers of English and the stimuli that are presented in this study were specifically chosen to represent items that contain sounds that are considered to be very difficult for non-native speakers to produce , such as, for example, the phonemes [v], and

word final [b] (Broersma & Kolkman, 2004). For a complete overview of these items, see Appendix B.

This study also considers the effects of differences in exposure to the target language by investigating second language acquisition in The Netherlands, a setting in which the second language, in this case English, is not an official language of communication. With the recent introduction and national recognition and regulation of English-Dutch bilingual education programs in secondary schools, there now are two distinct populations of English language learners who receive similar exposure to the target language outside of the educational setting, but who are distinctly different in the amount of instruction they receive in the target language. The main difference between the exposure of the two non-native speaker subject groups in this study to the target language, thus, lies in the amount to which subjects are exposed to target language instruction and, as such, allows for the variable 'exposure to the target language' to be considered as a factor in second language acquisition alongside a possible critical period for second language acquisition.

2.1.4 Bilingual Education

Independent of the situation in The Netherlands, the degree to which students are exposed to a target language in their education can be considered a controversial matter. Next to monolingual education, or the type of education in which students receive instruction in a single language, which often is their mother tongue, bilingual education, or that type of education in which students receive instruction in two languages, has become increasingly popular as a type of education. Ever since it's first implementation, however, the effectiveness of bilingual education has been a topic of discussion. In order to investigate the effectiveness of this type of education in comparison to that of monolingual education, it is important to carefully define what it consists of.

Although there are multiple definitions of bilingual education that are used by researchers when discussing a variety of situations and processes within the field of education, for the purpose of this thesis, bilingual education is defined along the lines of Genesee (1987) as a form of education in which a substantial part of the curriculum is taught in a language other than the first language of instruction, with the first language of instruction often being the native language of the majority of students. This definition, too, is rather general, but considering the existing variety of programs in bilingual education, a more narrow description seems difficult to establish. In order to classify the different types of bilingual education, a number of distinctions can be made, two of which are concerned with 'immersion', or the exposure of students to the target language:

1. *Late Immersion versus Early Immersion*, in which the difference is determined by the age at which a student initially starts to receive bilingual education. The term early immersion is generally used to refer to programs that start bilingual instruction around Kindergarten or first grade, whereas late immersion is mostly used for programs which do not commence until secondary level education.
2. *Full Immersion Programs versus Partial Immersion Programs*, in which the level of immersion in the target language determines the type of program. Partial Immersion programs are programs in which students receive a portion of their instruction in a target language, whereas full immersion programs provide instruction only in the target language. Technically, full immersion programs are therefore not types of bilingual education, though some researchers, such as Krashen (1997) for example, argue that if the language of instruction is not the native language of the student, this still counts as bilingual education. Within partial immersion programs, the degree of immersion may vary considerably, though, as Genesee (1987) mentions,

"...programs in which the second language is [only] used for teaching language arts..."(p. 19) are excluded.

In addition, distinctions can also be made based on the type of bilingualism that is involved:

1. *Elective Bilingualism versus Circumstantial Bilingualism* in which the term 'bilingualism' refers to the ability to communicate with native-like proficiency in two languages. "Elective bilingualism", according to Baker (2001) is " a characteristic of individuals who chose to learn a language" (p.3), whereas "circumstantial bilingualism" could be defined as a learner's attempt to " learn another language to survive" (p.3). Circumstantial bilingualism, thus, is imposed upon the learner by his or her environment, whereas elective bilingualism constitutes a choice.
2. *Additive Bilingualism versus Subtractive Bilingualism*, in which additive bilingualism refers to the process in which a second (or other) language is added to the language repertoire of the student, whereas subtractive bilingualism is used to describe the situation in which a second (or other) language has replaced the student's native tongue.

Essentially, regardless of the different types of bilingual education that exist, the main difference between any type of bilingual education and monolingual education lies in the medium through which students receive their instruction. In this thesis, a specific type of late, partial, elective and additive bilingual education in The Netherlands is compared with a monolingual type of education in order to investigate the effects of exposure on the attainment of native-like pronunciation in English.

2.2 Second Language Acquisition and The Netherlands

Before examining the proposed subject groups, it is important to consider the

situation in which their language acquisition takes place.

2.2.1 Situational Considerations

The Dutch language belongs to the Germanic branch of the Indo-European language family, and, as such, is closely related to the English language. During recent years, the English language has become an especially important source for language borrowing, mainly within the domains of technology and media such as television and the Internet. Due to this language borrowing, the linguistic gap between the English and the Dutch language has become increasingly smaller in terms of overlapping lexicon, but it has not yet been bridged entirely. Although both languages are linguistically similar, there are still considerable phonological differences that make it difficult for speakers of Dutch to obtain a native-like pronunciation of English. Not only are phonemes such as [v] and [θ] not native to the phoneme inventory of the Dutch language, specific rules for the pronunciation of sounds, such as word-final devoicing of voiced consonants such as [d] and [b] in Dutch for example, often result in a non-native accent of many Dutch speakers of English (Broersma & Kolkman, 2004).

As The Netherlands are situated fairly centrally in Europe, inhabitants of The Netherlands can easily immerse themselves in settings in which foreign languages are spoken. Although explicit foreign language instruction does not generally commence until the secondary level of Dutch education, which is comparable to American grade seven, students are increasingly exposed to international contacts outside of the instructional domain as they grow older. This international immersion can provide motivation for foreign language learning, if students are motivated to acquire a foreign language based on their desire to use this language to integrate in a target culture, when on vacation, for example.

Dutch learners of English can easily immerse themselves in English as the media provide a high degree of passive language exposure such as through television shows and movies which are often produced in the English language with Dutch subtitles. In addition, Internet websites are often written in English. Although television shows may be subtitled, Dutch speakers of English generally receive a considerable degree of passive exposure to the target language as they still receive auditory English language input, alongside Dutch translations.

According to Krashen's 'Input Hypothesis', comprehensible input, which is input that can be understood by the student and which is slightly above the students' mastered level of comprehension, (Krashen, 1981), is necessary for acquiring a second language. According to this hypothesis, passive exposure to English through the media could also be important to second language acquisition if the input were comprehensible. It is possible that Dutch students may have learned to distinguish between different English phonemes through the passive exposure before their secondary schooling has commenced.

2.2.2 The Monolingual Education System

The Dutch education system is mainly monolingual and can be divided into two types of schooling: primary and secondary schooling. Primary schooling takes place for a total of 8 years, including Kindergarten, and is arranged through a homogenous grouping of students by age. Parents can chose to enroll their children in any primary school they like, but no transportation to and from school is provided. During these 8 years of education, students are prepared for their secondary education and ongoing evaluation takes place in order to determine the level of the student. Upon completion of the eighth year of secondary schooling, the equivalent of the sixth grade in the United States, students are

required to take a final exam which, together with an 'advice' or evaluation by the student's primary instructor, determines the type of secondary education a student will receive.

Students then continue their education in one of the following three levels of secondary education: Voorbereidend Wetenschappelijk Onderwijs (VWO),which is the highest level of secondary education, Hoger Algemeen Voortgezet Onderwijs (HAVO), which is the middle level of secondary education, and Voorbereidend Middelbaar Beroeps Onderwijs (VMBO), which is the lowest level of secondary education. In secondary schooling, thus, students are generally grouped based on their level of achievement, though some secondary schools chose to continue to place students into homogenous groups for one or two additional years.

Unlike most tracking systems, however, the Dutch system is relatively flexible so that, if it becomes apparent that a student is placed in the wrong level, it is still possible for the student to get transferred. After the initial 2 years, it becomes more difficult to do so, however, due to the differing number of years of education each level consist of. As VWO programs take a total of 6 years, HAVO programs consist of 5 years of secondary education, and VMBO programs only take 4 years to complete, it is much more difficult for students in their third or fourth year of secondary education to transfer onto a different level than if such event were to take place during the first or second year of secondary schooling.

It is, generally, not until this secondary education, that students are exposed to foreign language learning in their educational careers. Although programs vary widely in their requirements, national regulations stipulate that students receive at least 4 years of instruction in at least two foreign languages. Depending on the level of education students receive, requirements on foreign language learning can be extended to require a minimum of five years of foreign language instruction in up to five foreign languages. English language

classes in monolingual secondary programs generally consist of a type foreign language instruction that integrates all four skills (reading, writing, listening, and writing), and focuses on content such as vocabulary, grammar, and culture of the target language. The exact curriculum varies per institution.

In addition to the primarily monolingual Dutch education system, a number of schools have, recently, chosen to develop bilingual programs in which their students may chose to enroll.

2.2.3 Bilingual Education Programs

Bilingual education is a fairly recent phenomenon in The Netherlands and was initially introduced in 1989 at international school Alberdingk Thijm in Hilversum when the school created a special English-Dutch bilingual division for Dutch students. As the school was already offering English medium education for international students, the only difference between their regular curriculum and their new program concerned the student population. In addition to teaching international students, they opened their school to native speakers of Dutch, who, with a shared first language, received instruction partially in English, and partially in their native language.

Although a small number of schools chose to gradually follow this example by adding a bilingual program to their otherwise monolingual curriculum, bilingual education did not grow into a nationally organized or recognized form of education until the late 1990s. With the national recognition of bilingual education, or Content and Language Integrated Learning (CLIL) as it is generally known as in The Netherlands, it became possible for bilingual programs to obtain government funding, so that, over time, the number of bilingual programs has rapidly increased. Content and Language Integrated Learning is a type of education in which content and language are integrated into one whole.

Similar to content-based instruction, CLIL generally involves teaching students language, such as English for example, through content, such as history for example. According to Landelijk Netwerk voor Tweetalig Onderwijs (January, 2009):

- there are 100 schools with a bilingual VWO -stream
- four schools are preparing for Content and Language Integrated Learning
- there is one school with CLIL in German
- these schools together have more than 15,000 CLIL-students

Although schools have considerable freedom in determining their curriculum, there are a number of basic requirements that need to be met by all bilingual programs as described in Landelijk Netwerk voor Tweetalig Onderwijs (2008):

1. A maximum of 50% of the lessons may be taught in English
2 CLIL should be financially self-supporting
3 CLIL cannot be detrimental to students' Dutch

In addition, if schools wish to obtain recognition and certification for their bilingual programs, they must adhere to the CLIL standard that was created by the European Platform for Bilingual Education and the Landelijk Netwerk voor Tweetalig Onderwijs.

Bilingual programs in The Netherlands consist of what Baker (2001) terms "late partial immersion education" (p.217) in which students start their bilingual education during their secondary education, and only receive a part of their instruction through English. Although programs vary considerably, instruction in the target language does not start until secondary education and a maximum of 50 % of all nationally required courses is taught in the target language.

In addition to consisting of a late and partial type of immersion, bilingual programs in The Netherlands aim to develop an elective type of bilingualism. Students choose to receive bilingual education through English and Dutch in order to become bilingual, which sets them apart from students in monolingual Dutch programs for whom English is a required course. As such, it could well be possible that students in both programs have

different motivations for obtaining the English language, which could influence their second language attainment. Although information on learners' motivation was collected through the use of a questionnaire and considerable differences between students in bilingual programs and those in monolingual programs were not observed, this variable has not statistically been controlled for in this study, and as such, may present a limitation. Researchers such as Huibregtse (2001), however, have investigated the difference in motivation between students in bilingual and monolingual secondary education programs in The Netherlands and found that it does not significantly influence the English language attainment of students at the VWO level.

Most secondary bilingual programs in The Netherlands are offered at the VWO level , though a small number of schools offer bilingual programs at HAVO or at VMBO levels as well. Due to ' The Second Phase', most schools, however, only have a bilingual track for secondary grade one through three as it is often economically impossible for schools to maintain bilingual education in higher grades as there are not always enough bilingual education students in each track. As a solution, some schools offer special English classes for grades four to six, which may be counted as elective courses, during which students are immersed in the target language and can be prepared for specific language tests, therewith providing them with the opportunity to earn special certificates or degrees, such as the International Baccalaureate degree in addition to their nationally recognized diploma.

As bilingual education is still a relatively new phenomenon in The Netherlands, not much research has been dedicated to the topic yet. Although some studies have focused on the pedagogical aspects of bilingual education, not many have researched the effectiveness of his type of education in The Netherlands. Huybregtse (2001) is one of the few studies that did focus on this matter as it investigated the effectiveness of bilingual education programs

in The Netherlands through an evaluation of the English language proficiency of students who had been enrolled in bilingual programs. Huybregtse (2001) found that bilingual programs increase the second language proficiency of Dutch students, although this study does not specify the degree to which this is the case. As well, in her methods, Huybregtse focused solely on reading exams and vocabulary tests, so that her findings cannot necessarily be extended to other skill components of second language acquisition. In order to add to this research, Pantophlet (2008) investigated both the results of students in bilingual and monolingual program on the reading and listening components of the national final examinations that students are required to take. This study, as well, found positive effects of increased exposure to the target language on students' second language acquisition, but did not take speaking or writing skills into consideration. Additionally, these studies did not test the CPH as no comparisons were made between non-native language learners and native speakers.

CHAPTER 3

METHODOLOGY

3.1 Subjects

In order to investigate the research questions mentioned above, two groups of subjects were included in this study. The first group of subjects consisted of a total of 18 subjects who provided three separate speech samples and, for this reason, will be referred to as 'speakers'. The second group consisted of 21 subjects who functioned as raters of speech samples and will be referred to as 'raters'.

The speaker group consisted of 6 native speakers of English and 12 native speakers of Dutch who are L2 speakers of English. Each subject recorded three different speaking tasks. The non-native English speaking subjects in this study had not received formal instruction in English until after the CPH for phonology is supposed to have passed, around age 12 (Long, 2007), and all subjects were in their final year of secondary education and at least 18 years of age. In order to minimize the effects of individual variables on the outcomes of this study, non-native speakers filled out a questionnaire containing questions on their language background and motivations, and only those non-native speakers who had spent their entire lives in The Netherlands and were raised monolingually in Dutch were included in this study. Native speakers, consequently, were only included in the study if they had lived their entire lives in a country in which English was the main language of communication and if they had been raised monolingually in English.

The non-native English speaking speaker group was further divided into two groups of 6 speakers depending on the type of secondary education they had experienced. The first group consisted of 6 students who received monolingual Dutch secondary instruction with an average of 2 hours of EFL instruction per week for a total of 6 years, and the second group consisted of 6 students who had received bilingual English-Dutch secondary education with an average of 2 hours of EFL and 18 hours of CLIL instruction per week for a total of 6 years.

Both groups, thus, have a similar history of traditional EFL instruction with a focus on language structure and the four skills, but bilingual subjects also received CLIL instruction, in other words content-based instruction through the English language. The rater group included in this study, on the other hand, consisted of 21 linguistically naïve native speakers of English who have lived their entire lives in a country in which English is the main language of communication. Subjects included in the rater group were all college-level students in the United States with no formal knowledge of linguistics and none have ever studied Dutch as a foreign language.

The subjects who were included in this study can be summarized as presented in Table 3.1.

Table 3.1 Distribution of Subjects Per Group

Type	Speakers	Raters
Native Speaker of English	6	21
Non-Native Speaker Monolingual Dutch	6	
Non-Native Speaker Bilingual English- Dutch	6	
Total:	18	21

3.2 Data Collection

As mentioned above, 'speaker' subjects were presented with three separate stimuli to which they were asked to respond, thus resulting in a total of 18 x 3 = 54 tokens of speech. Similar to Bongaerts et al. (1995), the tasks were presented to speakers in an order of increased speaker control and samples of the following types of speech were collected for each subject:

1. a word list
2. a paragraph
3. spontaneous speech

(See Appendix B for the actual speech elicitation stimuli that were used.)

Subjects were initially asked to read a word list out loud as though they were reciting a shopping list to a parent over the phone. The word-list that was used consisted of ten items and included a total of 54 instances of phonemes that are considered to be difficult for Dutch speakers of English to produce (Broersma & Kolkman, 2004). These phonemes are marked in Appendix B as well. An advantage of using a word list to elicit data, is that the content of responses will be similar across speakers, so that comparisons can be made. In addition, analysis of this type of speech considers the pronunciation of sounds alone, as supragsegmentals, such as stress and intonation, are not included. At the same time, this could also be a disadvantage as speaker control is very low and comparisons between speakers can only be made based on the pronunciation of sounds, therewith excluding the influence of suprasegmentals on second language pronunciation.

In order to allow for the influence of suprasegementals on second language pronunciation to be considered, subjects were then asked to read a paragraph from a novel out loud as though they were reading it to a close friend. This task involved the pronunciation of vocabulary as well as such matters as intonation and stress. This type of

task, however, still did not provide a high degree of speaker control as speakers were presented with a written version of what they were supposed to produce.

Finally, then, in order to allow for a higher degree of speaker control, subjects were asked to respond to an open-ended question aimed at eliciting spontaneous speech. As such, speakers had a high degree of control over the content of their response, which allows for matters such as word choice and register to be considered. Due to its high amount of speaker control, however, this task type also involves more domains of language than just pronunciation, as speakers control their use of word order and word choice as well. As such, making comparisons between speakers does not just involve detecting differences in their pronunciation, but it includes differences in other linguistic domains which may influence language proficiency as well.

As subjects were thought to be likely to respond somewhat differently to the final task, a decision was made to record a fixed amount of time for each response in order to manage data collection. Although Scovel (1981) demonstrates that native speaker raters can provide reliable judgments on the native-like pronunciation of their language within as little as eight seconds, and Bongaerts et al. (1995) successfully work with 16 to 20 second passages of recorded speech, for the purpose of this thesis, in order to make certain speech passages were long enough for native speaker raters to form a reliable opinion on their accentedness, a 30-second passage was recorded for each speaker.

After these speech samples were obtained, raters were asked to assign each sample a score on a 5-point Likert scale, identical to the one used in Bongaerts et al. (1995) based on the level of native likeness of each sample. An example of the rater evaluation sheet can be found in Appendix C.

Based on this scale, the following scores were assigned:

1. Very strong foreign accent; definitely non-native
2. Strong foreign accent
3. Noticeable foreign accent
4. Slight foreign accent
5. No foreign accent at all; definitely native speaker

Raters were told that there was a combination of native and non-native speakers who were asked to complete the assignment, though exact numbers of each type of speaker were not given. Raters did not receive any rater-training, because the purpose of this study was to investigate native speaker laypersons' perceptions of non-native speaker speech. However, each rater was presented with an anticipatory set of three example stimuli, in order to enable them to anticipate the range of possible levels of pronunciation they were about to encounter. Raters were not given any information as to whether the example stimuli were recorded by native or non-native speakers as this was thought to bare on the potential for influencing the evaluation of the stimuli with which they were presented.

3.3 Methods of Analysis

The 21 raters who took part in this study were randomly assigned to one of three groups, so that each rater group consisted of seven raters. Each rater group was then presented with a sequence of 18 stimuli, which included six samples of task one, six samples of task two, and six samples of task three and one token of each speaker. As such, each rater evaluated one token from each speaker, and each token from each speaker was rated by seven raters. Please see Appendix D for the chart that was used to sequence and organize the stimuli that were rated by each rater group.

In order to analyze the data that were obtained, mean scores were calculated for each subject and for each group. Standard deviations were also calculated for each group so as to mark the range in speaker pronunciation as rated by the raters in this study. Based on the

advantages and disadvantages of the three task types that were mentioned above, each task was analyzed separately in order to investigate its influence on second language proficiency.

Similar to the methodology used in Bongaerts et al. (1995), and Birdsong (1999), the mean score of each non-native speaker was then compared to the scores of the native speaker group for each task, and if this mean score fell within two standard deviations of the mean of the native speaker group, a non-native subject was considered to have a native-like pronunciation of English. Besides its history of being widely used and rarely criticized in similar studies on native-like pronunciation of second languages by non-native speakers, this measure was chosen to determine the native likeness of each subject because it captures the idea that non-native speakers can be considered to a pronounce language with a native-like ability if this speaker pronounces the language within the range of native speakers of the language as judged by an impartial, different set of native speakers. If a non-native speaker is, thus, rated by a native speaker rater similar to other native speaker speakers, a non-native speaker should be considered to have the capacity to speak the target language with a native-like fluency and such quality can carefully be measured by using normal distribution figures in which the variety in native speaker pronunciation, as observed by other native speakers, can be clearly marked for reference to that of non-native speakers. As this study aims to investigate whether it is possible for non-native speakers to obtain a native-like proficiency in the target language as judged by native speakers of that language, the two standard deviation criteria seems to provide a suitable way to analyze the speech of the subjects in this study. In using this measure, the total number of non-native speakers with a native-like language proficiency could be calculated so as to determine whether it is possible for the Dutch speakers in this study to obtain a native-like pronunciation of the English language as judged by native speakers of said language.

In addition, a mean comparison was made between the monolingual group and the bilingual group in order to see whether students in bilingual programs performed significantly better in speaking English than students in monolingual programs as based on native-speaker judgments for each task type. In order to evaluate the comparison that was made, one-tailed tests of significance (*t* test) were applied to the data. The alpha decision level was set at $p < 0.05$ before any calculations were made.

CHAPTER 4

RESULTS

In order to facilitate inter-group comparison, mean scores were calculated for individual subjects and subject groups per task. For a complete overview of scores that were assigned to each sample of each speaker by each rater, see Appendix E. As can be seen in Table 4.1, none of the non-native speakers' scores fell within two standard deviation from the mean of the native speaker group.

Table 4.1: Mean Scores Per Speaker on Task 1

Speaker	**Mean Score**
NS1	4.86
NS2	5.00
NS3	3.71
NS4	5.00
NS5	4.71
NS6	4.85
Group Mean	**4.69 (0.46)**
NNSM1	1.86
NNSM2	2.71
NNSM3	2.29
NNSM4	2.57
NNSM5	2.43
NNSM6	2.15
Group Mean	**2.34 (0.31)**
NNSB1	2.57
NNSB2	3.43
NNSB3	2.29
NNSB4	2.86
NNSB5	2.86
NNSB6	2.86
Group Mean	**2.81(0.38)**

In order to answer the second research question that was posed in Chapter 1, a *t* test was done in order to test the significance of the difference in means between the monolingual non-native group and the bilingual non-native group. Using the mean scores that can be found in Table 4.1, the observed value for *t* was found to be 2.374. As this observed score for *t* is larger than the critical value for *t*, the difference in means is significant.

Secondly, mean scores were calculated for all speaker groups based on their performance on the second task. As can be seen in Table 4.2, similar to the first task, none of the non-native speakers' scores fell within two standard deviation from the mean of the native speaker group for the second task. There were, thus, no non-native speakers who were considered to have a native-like pronunciation of English.

Table 4.2: Mean Scores Per Speaker on Task 2

Speaker	**Mean Score**
NS1	4.86
NS2	4.86
NS3	4.29
NS4	5.00
NS5	4.86
NS6	4.86
Group Mean	**4.79 (0.25)**
NNSM1	1.86
NNSM2	2.15
NNSM3	2.43
NNSM4	2.43
NNSM5	3.00
NNSM6	1.57
Group Mean	**2.24(0.50)**
NNSB1	2.43
NNSB2	3.14
NNSB3	2.57
NNSB4	2.57
NNSB5	3.71
NNSB6	2.15
Group Mean	**2.76(0.57)**

In addition, a *t* test was done in order to test the significance of the difference in means between the monolingual non-native group and the bilingual non-native group. Using the mean scores that can be found in Table 4.2, the observed value for *t* was found to be 1.677. As the observed score for *t* is smaller than the critical value for *t*, the difference in results between the monolingual and bilingual speaker groups for this task type are not significant.

After this, mean scores were calculated for all speakers and speaker groups based on their performance on the third task. As can be seen in Table 4.3, similar to the first and second task, none of the non-native speakers' scores fell within two standard deviation from the mean of the native speaker group for the third task. There were, thus, no non-native speakers who were considered to have a native-like pronunciation of English.

Table 4.3: Mean Scores Per Speaker on Task 3

Speaker	**Mean Score**
NS1	4.86
NS2	5.00
NS3	4.29
NS4	5.00
NS5	4.57
NS6	5.00
Group Mean	**4.79 (0.30)**
NNSM1	2.14
NNSM2	2.00
NNSM3	3.14
NNSM4	2.71
NNSM5	2.57
NNSM6	2.71
Group Mean	**2.55 (0.42)**
NNSB1	3.29
NNSB2	3.14
NNSB3	3.14
NNSB4	2.57
NNSB5	3.71
NNSB6	3.00
Group Mean	**3.14 (0.37)**

In addition, a *t* test was done in order to test the significance of the difference in means between the monolingual non-native group and the bilingual non-native group. Using the mean scores that can be found in Table 4.3, the observed value for *t* was found to be 2.592. As this observed score for *t* is, larger than the critical value for *t*, the chance is smaller than one per cent that the difference in means between the two groups is attributable to chance.

Finally, then, overall mean scores were calculated for each speaker and for each speaker group. As can be seen in Table 4.4, native speakers were overall scored higher than any of the non-native speakers and the students in bilingual programs scored as high as or higher than students in monolingual programs in terms of overall scores.

Table 4.4. Mean Scores per Speaker Overall

Speaker	**Mean Score**
NS1	4.85
NS2	4.95
NS3	4.10
NS4	5.00
NS5	4.71
NS6	4.90
Group Mean	**4.75 (0.33)**
NNSM1	1.95
NNSM2	2.29
NNSM3	2.52
NNSM4	2.57
NNSM5	2.67
NNSM6	2.14
Group Mean	**2.36 (0.35)**
NNSB1	2.76
NNSB2	3.24
NNSB3	2.67
NNSB4	2.67
NNSB5	3.43
NNSB6	2.67
Group Mean	**2.90 (0.35)**

In order to answer the second research question that was posed in Chapter 1, a *t* test was done in order to test the significance of the difference in means between the monolingual non-native group and the bilingual non-native group. Using the mean scores that can be found in Table 4.4, the observed value for *t* was found to be 2.673. As this observed score for *t* is, larger than the critical value for *t*, the chance is smaller than one per cent that the difference in means between the two groups is attributable to chance, which means that the difference in means is significant.

In addition to the statistics mentioned above, Crohnbach's alpha was used in order to calculate interrater reliability. For the three rater groups, the alpha's that were found are respectively 0.907, 0.951, and 0.948 which are generally considered to represent a high degree of interrater reliability.

CHAPTER 5

DISCUSSION

Based on the results as presented in the previous chapter, there were no non-native speakers of English whose scores fell within two standard deviations of the mean of the native speaker group for any of the three task types. In other words, there were no non-native speakers who were considered to have a native-like proficiency of English as judged by native speakers. The non-native speaker who obtained the highest rating on any of the separate tasks is NNSB5, with a score of 3.71. In order to qualify as a native speaker, however, a subject would have to have obtained a mean score of either 4.29 on the second task or at least 4.19 on the third task. As such, though NNSB5 was rated, on average, as having in between a noticeable foreign accent and a slight foreign accent, their pronunciation was not considered to be of similar quality to that of a native speaker.

Although Lenneberg (1967) had no access to such data when he proposed the CPH, it seems that, in fact, his perceptions were also accurate for L2 native-like pronunciation according to the data obtained in this study.

Although the lack of non-native speakers who have achieved a native-like pronunciation of English seems to provide support for a critical period for second language pronunciation from an EFL perspective, we must also realize that the results are not evidence of ultimate attainment, nor are they statements about the final native-like

pronunciation that these subjects might acquire or the effects of what additional exposure to the target language in varying degrees of intensity might produce. As such, the results merely indicate that, at this point, after 6 years of English language instruction, the subjects in this study have not been able to obtain a native-like pronunciation based on native speaker laypersons' perceptions. The English pronunciation of the subjects in this study may continue to evolve over time toward target language norms; therefore, this study does not directly provide support for a critical period for second language pronunciation by itself.

The non-native speakers who received bilingual education overall were rated closer to target language norms than the non-native speakers who received monolingual education. As such, subjects in the bilingual group were considered to have a more native-like pronunciation of English than subjects in the monolingual group. Is the reason for the bilingual group's advantage motivated by the amount of exposure to the target language over time, the type of exposure to the target language over time (i.e., the bilingual group received CLIL instruction), or a combination of both factors? This is a topic for further research in this area.

Although there were a number of instances in which speakers in monolingual programs received higher scores than particular speakers in bilingual programs on individual tasks, a *t* test of significance was applied to each task type in order to investigate whether the difference in mean scores between the non-native speaker groups was significant. As it turns out, the difference in mean scores between the two non-native speaker groups was indeed significant for both the first task and the third task. Although the difference in mean scores between the two non-native speaker groups was not significant for the second task, significance was found for the overall mean scores of the monolingual and bilingual non-native speaker groups.

The finding that students in bilingual programs in The Netherlands have a more native-like pronunciation of English than students in monolingual programs in The Netherlands supports the idea that the 'amount of input' learners receive plays an important role in obtaining a native-like proficiency. It may also indicate that type of input (i.e., traditional language instruction versus CLIL) plays an important role in as well. In CLIL classrooms, L2 learners are negotiating for content and languages with their peers and instructors. As such, the results presented in the previous chapter suggest that second language acquisition is influenced by both the intensity and type of input. This notion supports the idea that second language acquisition is not solely an innate process, but that there has to at least be an interaction between processes inside the learner and those external to it.

Interestingly, both non-native speaker groups received higher mean scores for the third task, which involved a higher degree of speaker control. As such, it is possible that the influence of other language skills influences second language pronunciation. A reason for this could be that, with an increased speaker control, speakers are able to compensate for the non-native features of their speech by avoiding them. As speakers were required to produce particular sounds in the first task, and particular sounds and suprasegmentals in the second task, they were required to demonstrate their proficiency in these two areas. In the third task, however, speakers were able to use only those sounds and suprasegmentals that they wished to use, and, as such, may have left out particular sounds and suprasegementals. Whether speakers indeed omit language skills that they struggle with when given more control on their output was beyond the scope of this project, but would be an interesting topic for further research.

A last point that needs to be mentioned, here, is that all three groups had low standard deviation scores, which means that the distribution of scores within each group was relatively homogenous. Subjects in each group were, as well, scored very similarly by all raters. The raters, thus, were able to distinguish between three different types of speaking proficiency. The individual scores that speakers received suggest that native-speaker raters were able to clearly distinguish between native speakers and non-native speakers with little variation among individual raters, even without rater training. Although the difference in mean scores between the two non-native speaker groups is smaller than when the NNS groups are compared to those of native speakers, the data do suggest that raters were able to distinguish among the three different groups, (i.e., NS, monolingual NNS, and bilingual NNS), and rated them differently. This difference emerged from the data, not from rater training. The almost identical standard deviation scores for each of these three groups suggests that native speaker raters were able to clearly identify the subjects in each group as adhering to a level of pronunciation that marked that group as a whole and that they were able to make judgments about the pronunciation of these speakers that could likely be generalized to other native speakers.

CHAPTER 6

CONCLUSION

In order to add to the existing body of research available on the CPH, this thesis investigated the possibility for such a hypothesis to explain differences in second language pronunciation from an English as a foreign language perspective. This study did not make use of preidentified advanced speakers of English, but focused on an opportunity sample of students from both monolingual and bilingual programs in The Netherlands.

In addition, the aim of this thesis was to investigate the influence of the amount of target language exposure on the acquisition of target language pronunciation skills. By including subjects who do not reside in a geographical location in which the target language is used as an official language of communication, the variable 'exposure to the target language' could be investigated in more precision, based on the type of education subjects had received, than if they had had ample opportunities to be immersed in the target language outside the instructional domain. Although subjects may receive some passive exposure to the target language outside of the instructional domain, such exposure is similar for all subjects and various studies, such as Kuppens (2007), have shown that the effects of such exposure are minimal on learners who speak Dutch or Flemish as a first language. Passive exposure through, for example, the media, prior to the critical period cannot result in a native-like ability to speak a target language

In order to structure this research, the following research questions were posed in Chapter 1:

1. Are Dutch L1 speakers who are late learners of English able to achieve a native-like pronunciation of English as based on comparison scores between samples of their speaking and those of L1 speakers of English?
2. Does exposure to the target language make a difference in the level of attainment of pronunciation in Dutch L1 speakers who are late learners of English? In other words, is there a difference in English language pronunciation between students in monolingual education programs and those in bilingual education programs in The Netherlands as based on comparison scores between samples of their speaking?

Based on the results that were presented in Chapter 4, these questions can be answered in the following manner:

1. Dutch L1 speakers who are late learners of English are not necessarily able to achieve a native-like pronunciation of English as based on comparison NS rater scores between samples of their speaking and those of L1 speakers of English.
2. Exposure to the target language makes a difference in the level of attainment of pronunciation in Dutch L1 speakers who are late learners of English. In other words, there is a difference in English language pronunciation between students in monolingual education programs and those in bilingual education programs in The Netherlands as based on comparison scores between samples of their speaking.

Research question one, thus, warrants acceptance of the H0 hypothesis, which states that there were no non-native speakers of English who were considered to have obtained a native-like pronunciation of English by native speaker raters.

The fact that there are no non-native speakers of English in this study who were considered to have obtained a native-like proficiency in the English language, seems to provide support for a critical period for second language pronunciation from an EFL perspective, but the results do not represent ultimate attainment, or the final proficiency someone can acquire. More longitudinal research is needed to establish the effects of a critical period for second language pronunciation on ultimate attainment, possibly through replicating the current study and keeping track of the development of subjects' language attainment for a period of time. It would be important in such study not to use preidentified advanced speakers of a target language, but rather to use opportunity samples to investigate the ultimate attainment of target language learners with a range of language abilities in order to get a clearer picture of the different factors involved second language pronunciation.

In addition, it would be beneficial, if possible, to include subjects who have only been exposed to one particular variety or dialect of English during their English language instruction, so as to minimize the effects of interactions between different varieties of English on the pronunciation of non-native speakers of English. Although the data suggest that the raters in this study did not seem to have considerable difficulty in classifying speakers based on their pronunciation, it is possible that features from varieties of English other than American English in speakers pronunciation may have influenced raters' judgments. As such, this could be considered a limitation to this study, which in future research should be avoided or minimized, as far as possible.

The second research question posed in Chapter 1 requires the acceptance of the H1 hypothesis for that question, which states that subjects who received English- Dutch bilingual education in The Netherlands had obtained a more native-like pronunciation of English than subjects in monolingual Dutch programs in The Netherlands as judged by native speakers of English.

The finding that students who had received bilingual secondary education were, overall, rated higher than students who received monolingual secondary education, supports the idea that 'amount of input' plays a large role in obtaining a native-like pronunciation of English. In addition, this finding suggests that bilingual programs are a more successful way of increasing proficiency than monolingual programs in The Netherlands as bilingual programs provide both more varied and more intense input. This result implies, as well, that a maximized level of exposure to a target language in English as a Foreign Language settings is beneficial to the development of students second language pronunciation skills and is, therefore, an important factor to consider in designing second language curricula, both in monolingual and bilingual education programs.

In conclusion, although further research is needed in order to ascertain the effects of a critical period for second language pronunciation on ultimate attainment, the amount to which learners are exposed to a target language in the instructional domain significantly influences their native-like pronunciation of that language. As such, this thesis adds to the body of research available on second language acquisition by investigating the effects of exposure the target language as a factor in the development of second language pronunciation and by considering the CPH from an English as a Foreign Language perspective.

APPENDIX A

QUESTIONNAIRE

Please answer the following questions about your personal characteristics and the circumstances under which you have learnt the English language. This information is anonymous and will only be used to provide the researcher with information that can influence the reliability of the results of the research study.

1. Gender:.....

2. Age:.....

3. Age at which you formally started learning English:.....

4. Do any of your parents speak English as a first language?.....

5. Number of years of formal instruction in English:.....

6. Type of program in which you learnt English (e.g., Monolingual, Bilingual) :.....

7. Average number of hours you received instruction in English per week:.....

8. Have you ever lived in an English speaking country?.....

If so, for how long?.....

9. Is there any other way in which you have been exposed to the English language(media, travel etc.) ?.....

.....

10. Why do you study English?
a). because I have to
b.) because I like the language
c.) because my parents want me to
d). Other:.....

11. On a scale from 1-5 with 5 being the highest and 1 being the lowest, how would you rate your motivation to learn English? .....

APPENDIX B

SPEECH ELICITATION MATERIALS

Word List

Your mother has asked you to get some groceries at the store, but the paper on which you have listed the items that you need has gotten wet from the rain outside and is very hard to read. You call your mother in order to make sure you will pick up everything she needs. Below is the list of items that you think are written on the wet piece of paper, repeat these to your mother so that she can tell you whether the items are indeed the ones she needs and whether there is anything missing.

English sounds that are difficult to pronounce for Dutch L1 speakers:

consonants: [th],[□,r] (final)devoicing: [v,f], [d,t], [z,s], [b,p], [g,k], (35)
vowels: [ae, □] [□:, □, a:] [□, □] (20)

1. A bottle of Bathtub cleaner

2. An ounce of Cheddar cheese

3. A bag of frozen peas

4. A box of rat poison

5. Three packs of flashcards

6. A pound of prime rib

7.Two jars of walnuts

8. Twelve red roses

9. Two tins of cat food

10. A box of fruit leather

Short Paragraph

You have just finished reading a book and are very excited about it. Because you like the book so much, you want to tell your best friend about it, so that they might read the book as well.

Read the following paragraph out loud to your friend in order to show him or her what the book is about.

> Afterward, I find Morrie Schwartz, my favorite professor, and introduce him to my parents. He is a small man who takes small steps, as if a strong wind could, at any time, whisk him up into the clouds. In his graduation day robe, he looks like a cross between a biblical prophet and a Christmas elf. He has sparkling blue-green eyes, thinning silver hair that spills onto his forehead, big ears, a triangular nose, and tufts of graying eyebrows. Although his teeth are crooked and his lower ones are slanted back--as if someone had once punched them in--when he smiles it's as if you'd just told him the first joke on earth. He tells my parents how I took every class he taught. He tells them, " You have a special boy here." Embarrassed, I look at my feet. Before we leave, I hand my professor a present, a tan briefcase with his initials on the front. I bought this the day before at a shopping mall. I didn't want to forget him. Maybe I didn't want him to forget me.
>
> (excerpt from Mitch Albom's *Tuesdays with Morrie,* p. 4)

Spontaneous Speech

You and your friend are planning on hanging out over the weekend. You think about going to the movies and you try to find out which kinds of movies both of you like. Your friend tells you about his favourite movie and then asks you what your favourite (English) movie is. Give your friend a brief description of your favourite (English) movie and tell him why you liked it so much.

APPENDIX C

RATER EVALUATION SHEET

You are about to hear 18 sound fragments that were recorded by both native and non-native speakers of English. For each fragment, decide how native-like this sounds to you. You can give the following scores for each fragment:

1 Very strong foreign accent; definitely non-native
2 Strong foreign accent
3 Noticeable foreign accent
4 Slight foreign accent
5 No foreign accent at all; Definitely native speaker

Mark each sample by putting an 'x' in the column that corresponds to the score you would like to assign.

Sample/ Score	1	2	3	4	5
1.					
2.					
3.					
4.					
5.					
6.					
7.					
8.					
9.					
10.					
11.					
12.					
13.					
14.					
15.					
16.					
17.					

APPENDIX D

STIMULI EVALUATION PER RATER GROUP

Table D.1: Stimuli Evaluation Per Rater Group

Stimuli	Rater Group 1	Rater Group 2	Rater Group 3
1.	NS1-1	NS3-1	NS5-1
2.	NNSM1-1	NNSM3-1	NNSM5-1
3.	NNSB1-1	NNSB3-1	NNSB5-1
4.	NS2-1	NS4-1	NS6-1
5.	NNSM2-1	NNSM4-1	NNSM6-1
6.	NNSB2-1	NNSB4-1	NNSB6-1
7.	NS3-2	NS5-2	NS1-2
8.	NNSM3-2	NNSM5-2	NNSM1-2
9.	NNSB3-2	NNSB5-2	NNSB1-2
10.	NS4-2	NS6-2	NS2-2
11.	NNSM4-2	NNSM6-2	NNSM2-2
12.	NNSB4-2	NNSB6-2	NNSB2-2
13.	NS5-3	NS1-3	NS3-3
14.	NNSM5-3	NNSM1-3	NNSM3-3
15.	NNSB5-3	NNSB1-3	NNSB3-3
16.	NS6-3	NS2-3	NS4-3
17.	NNSM6-3	NNSM2-3	NNSM4-3
18.	NNSB6-3	NNSB2-3	NNSB4-3

NS: Native Speaker
NNSM: Non-Native Speaker Monolingual
NNSB: Non-Native Speaker Bilingual

APPENDIX E

SPEAKER SCORES PER SAMPLE

Table E.1: Speaker Scores Per Sample Rater 1-7

Sample	**NS11**	**NNSM11**	**NNSB11**	**NS21**	**NNSM21**	**NNSB21**	**NS32**	**NNSM32**	**NNSB32**
Rater									
1	4	1	1	5	1	4	4	2	3
2	5	2	2	5	2	2	4	1	1
3	5	2	2	5	1	3	4	1	1
4	5	1	3	5	4	3	5	4	4
5	5	3	4	5	3	4	5	4	3
6	5	2	3	5	5	4	4	2	3
7	5	2	3	5	3	4	4	3	3
Total:	34	13	18	35	19	24	30	17	18
Mean:	4.86	1.86	2.57	5	2.71	3.43	4.29	2.43	2.57

NS42	**NNSM42**	**NNSB42**	**NS53**	**NNSM53**	**NNSB53**	**NS63**	**NNSM63**	**NNSB63**
5	1	2	5	1	4	5	1	3
5	2	3	5	2	4	5	5	4
5	2	2	2	5	4	5	3	4
5	2	1	5	1	2	5	2	1
5	4	4	5	3	4	5	3	3
5	3	3	5	3	4	5	2	3
5	3	3	5	3	4	5	3	3
35	17	18	32	18	26	35	19	21
5	2.43	2.57	4.57	2.57	3.71	5	2.71	3

Table E.2: Speaker Scores Per Sample Rater 8-14

Sample	NS31	NNSM31	NNSB31	NS41	NNSM41	NNSB41	NS52	NNSM52	NNSB52
Rater									
8	4	3	2	5	3	3	5	4	4
9	5	3	1	5	3	4	5	3	4
10	4	2	3	5	2	3	5	3	4
11	4	3	3	5	3	2	5	3	4
12	3	2	4	5	2	4	5	3	4
13	3	2	1	5	3	2	5	2	3
14	3	1	2	5	2	2	4	3	3
Total:	26	16	16	35	18	20	34	21	26
Mean:	3.71	2.29	2.29	5	2.57	2.86	4.86	3	3.71

NS62	NNSM62	NNSB62	NS13	NNSM13	NNSB13	NS23	NNSM23	NNSB23
5	2	3	5	4	4	5	2	3
5	3	2	5	2	4	5	1	3
5	1	3	5	2	2	5	2	2
5	1	2	5	2	3	5	1	3
5	1	1	5	1	4	5	5	3
4	2	2	5	2	3	5	1	4
5	1	2	4	2	3	5	2	4
34	11	15	34	15	23	35	14	22
4.86	1.57	2.15	4.86	2.14	3.29	5	2	3.14

Table E.3 Speaker Scores Per Sample Rater 15-21

Sample	NS51	NNSM51	NNSB51	NS61	NNSM61	NNSB61	NS12	NNSM12	NNSB12
Rater									
15	5	3	4	5	3	3	5	3	4
16	5	2	4	5	3	4	5	2	4
17	5	3	3	5	2	3	5	2	2
18	4	1	1	5	1	2	4	1	2
19	5	4	3	5	1	2	5	1	1
20	4	1	1	4	2	3	5	2	1
21	5	3	4	5	3	3	5	2	3
Total:	33	17	20	34	15	20	34	13	17
Mean:	4.71	2.43	2.86	4.86	2.15	2.86	4.86	1.86	2.43

NS22	NNSM22	NNSB22	NS33	NNSM33	NNSB33	NS43	NNSM43	NNSB43
5	2	4	5	3	3	5	4	4
5	3	4	5	3	4	5	3	2
5	3	4	4	4	3	5	2	2
5	1	2	3	1	2	5	1	1
5	2	3	4	3	2	5	3	2
4	2	1	4	3	4	5	3	3
5	2	4	5	3	4	5	3	4
34	15	22	30	20	22	35	19	18
4.86	2.15	3.14	4.29	2.86	3.14	5	2.71	2.57

SELECTED BIBLIOGRAPHY

Baker, C. (2001). *Foundations of bilingual education and bilingualism.* Clevedon: Multilingual Matters.

Bialystok, E., & Hakuta, K. (1999). Confounded age: linguistic and cognitive factors in age differences for second language acquisition. In D. Birdong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161-181). Mahwah: Erlbaum.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68,* 706-755.

Birdsong, D. (1999). *Second language acquisition and the critical period hypothesis.* Mahwah: Erlbaum.

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. Kroll, and A. De Groot (Eds.), *Handbook of bilingualism* (pp. 109-127). New York: Oxford University Press.

Birdsong, D. & Paik, J. (2008). Second language acquisition and ultimate attainment. In B. Spolsky, and F. Hult (Eds.), *Handbook of educational linguistics* (pp. 424-436). London: Blackwell.

Bongaerts, T., Planken, B., and Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton, and Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30-50). Philadelphia: Multilingual Matters.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: the case of very advanced late L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133-159). Mahwah: Erlbaum.

Bongaerts, T. (2005). Introduction: Ultimate attainment and the critical period hypothesis for second language acquisition. *International Review of Applied Linguistics in Language Teaching, 43,* 259-267.

Broersma, M., and Kolkman, K. M. (Eds.). (2004). Proceedings of the 8th international conference on spoken language processing (Interspeech) 2004: *Lexical representation of non-native phonemes.* Korea: Jeju.

Brown, H. Douglas. (2007). *Teaching by principles: An interactive approach to language*

pedagogy. White Plains: Pearson Education Inc.

Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language, 63*, 544-573.

Cummins, J. and Swain, M. (1986). *Bilingualism in education: Aspects of theory, research and practice.* New York: Longman Inc.

Cummins, J. (1992). Language proficiency, bilingualism, and academic achievement. In P.J. Richard- Amato, and M.A. Snow (Eds.), *The multicultural classroom: Readings for content area teachers.* White Plains: Longman Publishers.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern day wild child.* New York: Academic Press.

DeKeyser, R. and Larson-Hall J. (2005). What does the critical period really mean? In J. Kroll, and A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp 89-108). Oxford: Oxford University Press.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Flege, J.E. (1987). A critical period for learning to pronounce foreign language? *Applied Linguistics, 8*, 162-177.

Flege, J. E., Munro, M.J., and MacKay, I.R.A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3138.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., and Tsukada, K. (2005). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics, 34*, 153-175.

Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education.* Rowley: Newbury House Publishers.

Grabe, W. and Stoller, F.L. (1997). Content-based instruction: Research foundations. In M.A. Snow, and D.M. Brinton (Eds.), *The content-based classroom* (pp. 5-21). White Plains: Longman Publishing Company.

Greene, J. (1998). *A meta-analysis of the effectiveness of bilingual education.* Claremont: Thomas Rivera Policy Institute.

Huibregtse, I., Admiraal, W., de Bot, K., Coleman, L., and Westhoff, G. (2000). Effecten en didactiek van tweetalig voortgezet onderwijs. *Toegepaste Taalwetenschap in Artikelen 64-2*, 9-20.

Huibregtse, I. (2001). *Effects and pedagogy of secondary bilingual education in The Netherlands.* Utrecht: Drukkerij Elinkwijk B.V.

Ioup, G., Boustagui, E., Tigi, M.E., and Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLC in a naturalistic environment. *Studies in Second Language Acquisition, 16*, 73-98.

Itard, J. (1962). *The wild boy of Aveyron.* (G. Humphrey & M. Humphrey, Trans.). New York: Appleton-Century-Crofts.

Krashen, S., Long, M., and Scarcella, R. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly, 13*, 573-582.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon Press.

Kuppens, A. (Dec, 2007). The effect of the use of English media on English vocabulary acquisition: An empirical study with Flemish youngsters. *Tijdschrift voor Communicatiewetenschappen, 34-4*, 325-336.

Landelijk Netwerk voor Tweetalig Onderwijs (2008). *Content and language integrated learning.* Retrieved from http://www.netwerktto.europeesplatform.nl

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Lengyel, Z. & Singleton, D. (1995). *The age factor in second language acquisition: A critical look at the critical period hypothesis.* Philadelphia: Multilingual Matters.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12,* 251-286.

Long, M. H. (2007). Chapter 3: Age differences and the sensitive periods controversy in SLA. In *Problems in SLA.* Mahwa: Erlbaum.

Marinova-Todd, S.F., Marshall, D.B., and Snow, C.E. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly, 34-1*, 9-34.

Mayberry, R. (1993). First language acquisition after childhood differs from second language acquisition; the case of American sign language. *Journal of Speech and Hearing Research, 36-6*, 1258-1270.

Pantophlet, D. (unpublished). Bilingual education in The Netherlands; a case study. *University of Amsterdam.* 2008.

Patkowski, Mark S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics, 11-1*, 73-89

Penfield, W. and Roberts, L. (1959). *Speech and brain mechanisms.* Princeton: Princeton University Press.

Pullvermüller, F. and Schumann, J. (1994). Neurobiological mechanisms of language

acquisition. *Language Learning, 44-4*, 681-734.

Schachter, J. (1988). Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics, 9*, 219-35.

Scovel, T. (1981). The effects of neurological age on nonprimary language acquisition. In R. Andersen (Ed.), *New dimensions in second language acquisition research* (pp. 33-42). Rowley: Newbury House.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics, 20*, 213-223.

Singleton, D. (2001). Age and second language acquisition. *Annual Review of Applied Linguistics, 21*, pp. 77-89.

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching, 43*, pp. 269-285.

Snow, C. and Hoefnagel-Hohle, M. (1978). The critical age for language acquisition: Evidence from second language learning. *Child Development, 49*, 1114-1128.

Swain, M., & Lapkin, S. (1982). *Evaluating bilingual education: A Canadian case study.* Clevedon: Multilingual Matters Ltd.

Swain, M.. (1993). The output hypothesis: just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*, 158–164.

Van Boxtel, S., Bongaerts, T., and Coppen, P.A. (2005). Native-like attainment of dummy subjects in Dutch and the role of the L1. *International Review of Applied Linguistics in Language Teaching, 43*, 355-380.

Westhoff, G. J. (1994a). *Effecten en didactiek van tweetalig voortgezet onderwijs.* Utrecht: WCC.

Westhoff, G. J. (1994b). *Tweetalig onderwijs in de praktijk.* Utrecht: WCC.

Woolfolk, A. (2007). *Educational psychology.* White Plains: Pearson Education Inc.